

Farklı Yöntemler Kullanılarak Geliştirilen Trafik Kaza Tahmin Modelleri ve Analizi

Ali Payıdar Akgüngör ve Erdem Doğan

Kırıkkale Üniversitesi Mühendislik Fakültesi İnşaat Mühendisliği Bölümü, Kırıkkale 71451, Türkiye
Tel: +90 318 357 42 42 Fax: +90 318 357 24 59 akgungor@kku.edu.tr, edogan@kku.edu.tr

Özet— Bu çalışmada, regresyon analizi, yapay sinir ağları (YSA) ve genetik algoritma (GA) yöntemleri kullanılarak İzmir ili için trafik kaza tahmin modelleri geliştirilmiştir. Modeller geliştirilirken nüfus, araç sayısı ve kaza sayısı model parametreleri olarak kullanılmış ve bu parametrelere ait 1986-2005 yılları arasındaki verilerden faydalanılmıştır. Regresyon analizi kullanılarak geliştirilen kaza modellerinde Smeed ve Andreassen kaza model formları kullanılmıştır. YSA modelinde 2-5-1 ağ mimarisi en uygun mimari olarak belirlenmiş, ağların gizli katmanında sigmoid, çıkış katmanında da doğrusal fonksiyon kullanılmıştır. Ağın eğitiminde ise ileri beslemeli geri yayılım algoritmasından yararlanılmıştır. GA tekniği ile modeller oluşturulurken farklı formlardaki modeller denenmiş ancak bu çalışma için en başarılı modelin üstel model olduğu görülmüştür. Geliştirilen bütün modellerin performansları ortalama mutlak yüzde hata (OMYH) ortalama mutlak hata (OMH) ve ortalama karesel hataların karekökü (OKHK) ölçütleri içinde değerlendirilmiştir.

Anahtar Kelimeler— Trafik kaza modeli, regresyon analizi, yapay sinir ağları, genetik algoritma

Abstract— In this study, traffic accident prediction models were developed using regression analysis, artificial neural networks (ANNs) and genetic algorithm (GA) methods for the city of İzmir. In the development of the models, population, the number of the vehicles and accidents were used as model parameters with data between 1986 and 2005. Smeed and Andreassen accident model structures were utilized in the development of the accident models by using regression analysis. In the ANN model, 2-5-1 network architecture was determined as the best suitable network architecture. The sigmoid and pureline functions were used as activation functions with feed forward-back propagation algorithm. In the GA approach, genetic algorithm models in different forms were developed but the exponential model form had the best performance for this study. The performances of all developed models were evaluated by the use of mean absolute percent errors (MAPE), mean absolute errors (MAE) and root mean square errors (RMSE).

Index Terms—Traffic accident model, regression analysis, artificial neural networks, genetic algorithm

I. GİRİŞ

Gelişmekte olan ülkelerde olduğu gibi ülkemizde de nüfus artışı ve ekonomik gelişmeler beraberinde araç sayısının hızla artmasına, araç sayısındaki bu artış ise karayolu ağı üzerindeki

taşıt hareketliliğinin büyümesine neden olmuştur. 1990 yılında 56 milyon olan ülke nüfusu yaklaşık 1.28 kat artarak 2008 yılı itibarı ile 72 milyona ulaşmıştır. Aynı dönem içindeki trafik göstergelerine bakıldığında araç sayısı 3.67 kat artarak 13 milyonu, araç-km değeri ise 2.70 kat artarak 72 milyarı geçmiştir. Görüldüğü üzere trafik değerlerindeki artış hızı, nüfus artış hızından daha fazla olmuştur. Araç sayısı ve trafik hareketlerindeki bu hızlı artışın bir sonucu olarak da ülkede meydana gelen trafik kazaları bugün ciddi boyutlara ulaşmıştır. Sadece 2008 yılında ülkemizde 929,304 trafik kazası meydana gelmiş ve bu kazalar neticesinde 4 bin den fazla kişi yaşamını yitirirken yaklaşık 184 bin kişide yaralanmıştır. Dolayısı ile 18 yıllık bir dönem içerisinde kazalardaki artış 8 katın üzerine çıkmıştır.

Yol güvenliği ile ilgili planlama ve politikaların belirlenmesinde, ileriye yönelik kaza tahmin modellerinden yararlanılmaktadır. Literatürde farklı yöntemlerle geliştirilmiş birçok kaza tahmin modeli bulunmaktadır. Bugün hala güncelliğini koruyan, en eski modellerden bir tanesi Smeed tarafından geliştirilen kaza modelidir. Smeed Kanunu olarak da bilinen bu modelde araştırmacı 1938 yılına ait 20 farklı ülkelerden aldığı veriler yardımıyla ölüm, araç sayısı ve nüfus arasındaki ilişkiyi incelemiştir [1]. Ancak Andreassen, Smeed modelinin verilerin bir yıla ait olduğunu, bir zaman serisi içermediğini ve modele ait sabit ve üstel değerlerinin her ülke için farklı olabileceğinden dolayı geliştirilen bu modelin, bütün ülkeler için uygulanamayacağını ifade ederek kendisi farklı bir model önermiştir [2]. Mekky gelişmekte olan ülkelerde taşıtlardaki hızlı artışın ölüm oranlarına olan etkisini araştırmış ve bu araştırmalarına bağlı olarak sanayileşmiş ülkeler ile gelişmekte olan ülkelerde araç başına ölüm oranları ile motorlu taşıt sayıları arasında ters bir ilişkinin olduğunu ortaya koymuştur [3]. Chakrobort ve Roy kaza büyüklük indeksi, kaza ölüm oranı, kaza ölüm riski ve kaza riskini kullanarak Hindistanın Kalkota şehrine ait yol güvenlik seviyesini belirlemeye çalışmışlar ve Smeed yaklaşımını kullanarak Kalkota şehri için kaza tahmin modelleri geliştirmişlerdir [4]. Valli Hindistan ve metropol kentleri için Smeed ve Andreassen bağıntularından faydalanarak kaza tahmin modelleri geliştirmiştir. Valli geliştirdiği modeller yardımıyla 2007 ve 2010 yılı için kaza, yaralı ve ölü sayılarını tahmin etmiştir [5]. Benzer şekilde Akgüngör ve Doğan Smeed ve Andreassen model yapılarından faydalanarak Türkiye için kaza, ölü ve yaralı sayılarını tahmin eden

modeller geliřtirmişler ve geliřtirdikleri bu modelleri üç farklı senaryo altında deęerlendirmişlerdir [6].

Son yıllarda yapay sinir aęları (YSA) da kaza modellemesinde ve karayolu güvenlięi ile ilgili çalıřmalarda yaygın olarak kullanılmaktadır. Mussone ve dięerleri İtalyanın Milan şehrindeki taşıtların sebep olduęu kazaları incelemek ve bir kaza modeli geliřtirmek için YSA teknięinden yararlanmışlardır [7]. Abdelwahap ve Abdel-Aty sinyalize kavřaklarda meydana gelen trafik kazalarında sürücülerdeki yaralanma řiddetini tahmin etmek için YSA teknięini kullanmışlardır [8]. Benzer řekilde Delen ve dięerleri çarpıřma ile ilgili faktörler ve yaralanma řiddeti arasındaki lineer olmayan iliřkiyi YSA yöntemi ile modellemeye çalıřmışlardır [9]. Chiou iki aracın çarpıřmasından kaynaklanan kazalar için YSA tabanlı bir uzman sistem kullanmışlar ve bir kazanın meydana gelmesinde yol geçiř hakkı, kaza yeri ve alkol kullanımının en etkili deęişkenler olduęunu tespit etmişler [10]. Akgüngör ve Doęan YSA ve lineer olmayan regresyon tekniklerini kullanarak Türkiye için kaza, yaralı ve ölü sayılarının tahmin etmişler ve farklı hata ölçütleri kullanarak her iki yöntemden elde edilen modelleri birleri ile karşılařtırmışlardır. Sonuçta YSA model tahminlerinin lineer olmayan regresyon modeline göre daha başarılı olduęunu ifade etmişlerdir [11].

Yapay zekânın dięer bir kolu olan Genetik Algoritma (GA), 1970'lerde John Holland tarafından ortaya atılmış ve bugün başta mühendislik problemleri olmak üzere birçok problemin çözümünde kullanılmaktadır. GA bir başlangıç popülasyonu olarak çeřitli seçimler yardımıyla en uygun çözüme yaklařmayı amaçlayan bir yöntemdir [12]. Ulařtırma problemlerinin çözümünde de bu yöntemden yararlanılmasına raęmen, kaza analizleri ve modellemelerinde bu yöntem yapılan sınırlı sayıda çalıřma bulunmaktadır. Haldenbilen ve Ceylan Türkiye için ulařtırma sektöründeki enerji talebini GA yardımıyla nüfus, gayri safi milli hâsıla, araç-km bağımsız deęişkenlerini kullanarak modellemeye çalıřmışlardır. Üç deęişik model formu (lineer, üstel ve kuadratik) için modellere ait ağırlık parametrelerini GA yardımıyla tespit ederek, 2020 yılına kadar çeřitli senaryolar dâhilinde enerji talebini tahmin etmişlerdir [13]. Gündoędu ve dięerleri, Erzurum ili örneğinde, trafik den kaynaklı gürültü seviyesi ile araç kompozisyonu arasında iliřki kuran modeller geliřtirmişlerdir. Modelleri geliřtirmek için GA yöntemini kullanan arařtırmacılar, girdi deęişkenleri olarak: araç sayımlarını, araç kompozisyonları için üst gürültü deęerlerini ve bölgedeki yüksek yapıların yol geniřliğine oranını kullanmışlardır. Oluřturulan modeller, mevcut ölçümler ve daha önceden geliřtirilmiş modeller ile karşılařtırıldığında nispeten iyi sonuçlar verdięini belirtmişlerdir [14]. Haldenbilen ve Ceylan, şehirlerarası ulařtırma talebini modellemek için GA yaklařımından yararlanarak lineer üstel ve kuadratik modeller oluřturmuşlardır. Ulařtırma talebini etkileyen üç ana parametre olarak nüfus, gelir ve araç sahiplięini kullanmışlardır. Geliřtirilen modelleri, en küçük kareler yöntemi ile

karşılařtırmışlar ve üstel model formunun dięer iki forma göre daha iyi sonuçlar ürettięini gözlemlemişlerdir. Senaryolar dâhilinde 2025 yılına kadar olan süreçte üstel model kullanılarak şehirlerarası ulařtırma talebini tahmin etmişlerdir [15]. Akgüngör ve Doęan yapay zekâ tekniklerinden olan YSA ve GA yaklařımlarını kullanarak Ankara iline ait kaza, yaralı ve ölü modelleri geliřtirmişlerdir. GA teknięi kullanılarak geliřtirilen modellerde Ankara ili için kaza ve ölü tahminlerinde üstel modelin, yaralı tahmininde ise lineer modelin hata ölçütleri içerisinde daha başarılı olduklarını ifade etmişlerdir [16].

Bu çalıřmada, nüfus ve motorlu araç sayılarına ait verilerden faydalanarak, İzmir iline ait trafik kaza sayıları regresyon analizi, YSA ve GA teknikleri kullanılarak tahmin edilmiştir. Aynı zamanda bu çalıřma ile üç farklı yöntemin bir karşılařtırması ortaya konarak, yöntemlerin birbirine olan üstünlükleri kaza modelleri örneğinde tespit edilmeye çalıřılmıştır.

II. REGRESYON ANALİZİ

Bir bağımlı deęişken ile bir veya daha fazla bağımsız deęişken arasındaki iliřkiyi incelemek amacı ile kullanılan istatistiksel bir yöntemdir. Bu yöntem yardımı ile deęişkenler arasındaki iliřki ortaya konarak bu iliřkinin modellenmesi yapılabildięi için başta mühendislik olmak üzere birçok alanda yaygın olarak kullanılmaktadır. Bu çalıřmanın bir bölümünde de İzmir iline ait kaza modelleri geliřtirilirken, lineer olmayan regresyon analizinden yararlanılmış ve regresyon modeli olarak da Smeed ve Andreassen kaza model formlarından faydalanılmıştır.

A. Smeed Modeli

Smeed 20 adet ülkenin nüfus ve araç sayıları ile ölü sayılarına ait 1938 yılı verileri kullanılarak denklem 1 de ifade edilen bir tahmin modelini geliřtirmiştir.

$$\frac{D}{N} = 0,0003 \left(\frac{N}{P} \right)^{-0,67} \quad (1)$$

Burada, D ölü sayısını, N motorlu araç sayısını, P ise nüfusu ifade etmektedir.

Smeed modelini geliřtirirken ölü sayısını bağımlı, araç sayısını ve nüfusu ise bağımsız deęişken olarak belirlemiş ve regresyon analizi için ařağıda verilen formu kullanmıştır.

$$\text{LogD} = \text{LogC} + B_1 \text{LogN} + B_2 \text{LogP} \quad (2)$$

Burada C, B₁ ve B₂ regresyon analizi tarafından belirlenen model sabitlerdir.

Yapılan regresyon analizi sonucunda ise Denklem 3 de verilen modeli geliřtirmiştir.

$$D = 0,0000993N^{-0,3377}P^{0,7323} \quad (3)$$

Ancak Smeed, geliştirdiği bu modeli kullanım kolaylığı için basitleştirerek aşağıda ifade edilen şekilde düzenlemiştir.

$$D = 0,0003(NP^2)^{1/5} \quad (4)$$

Smeed Denklem 4 de verilen modeli cebirsel olarak düzenleyerek, Denklem 1 de ifade edilen şekline dönüştürmüş ve alternatif bir form olarak kullanılabilceğini ileri sürmüştür.

B. Andreassen Modeli

Başta Andreassen olmak üzere, bazı araştırmacılar Smeed'in geliştirdiği modeli birçok yönü ile eleştirmişlerdir. Andreassen'in eleştirileri, Smeed modelini oluşturan verilerin bir yıllık olması, ülkelerin içinde bulunduğu farklı sosyal ve ekonomik şartlardan dolayı tek bir modelin bütün ülkeler için uygulanabilir olamayacağı ve Smeed'in modelini geliştirdiği yıllarda bilgisayar teknolojisinin sınırlı olmasından dolayı verilerin analizinde hataların yapılmış olabileceği yönündedir. Bunlara ilave olarak Andreassen'in yaptığı temel eleştirilerden birisi de Smeed modelinin alternatif formunu (Denklem 1) oluştururken yaptığı cebirsel düzenlemedir. Andreassen, regresyon analizi yapıldıktan sonra model üzerinde cebirsel düzenlemelerin yapılması bağımlı ve bağımsız değişkenleri ve değerlerini değiştireceğinden dolayı yeni değişkenlere göre regresyon analizinin tekrarlanması gerektiğini savunmaktadır.

Andreassen, yukarıda yapılan açıklamalar doğrultusunda ölü sayılarını tahmin için araç sayısı ve nüfusa bağlı olarak aşağıdaki model formunu önermiştir. Yazar, Almanya, Fransa, İngiltere ve ABD vb. her ülke için Denklem 5 de belirtilen farklı B_1 ve B_2 üstel değerlerini bulmuştur.

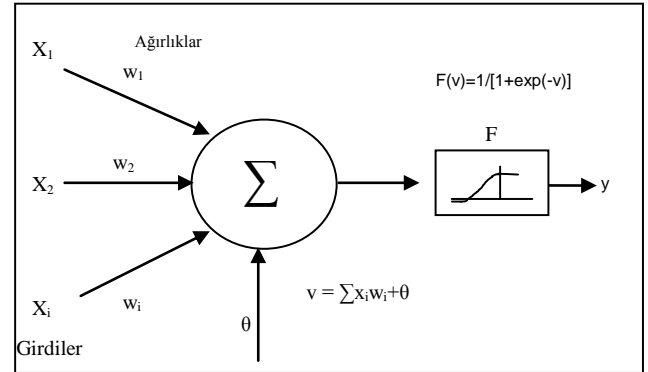
$$D = \text{sabit} * N^{B_1} P^{B_2} \quad (5)$$

III. YAPAY SİNİR AĞLARI

YSA, insan beyninin özelliklerinden esinlenerek geliştirilen bir yapıdır. Bu nedenle YSA, öğrenme, hatırlama, ezberleme, genelleme yapabilme ve bilgiler arasında ilişki kurabilme yeteneğine sahiptir. Beynin işlemlerini gerçekleştiren biyolojik nöronlar bulunmaktadır. Bu nöronlar kendilerine gelen sinyalleri bazı işlemlerden geçirerek daha sonraki nöronlara iletirler böylelikle bazı tepkiler oluşur. YSA'da da tıpkı beyinde bulunan nöronlar gibi yapay nöronlar bulunmaktadır. Bu yapay nöronların birbirlerine çeşitli mimarilerle bağlanması ile yapay sinir ağları oluşmaktadır.

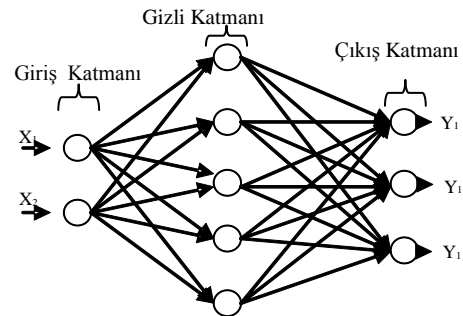
YSA'yı oluşturan temel işlemci yapay nöronudur. Bir yapay nöronda girdiler, ağırlıklar, toplama fonksiyonu, transfer fonksiyonu ve çıkış fonksiyonu bulunmaktadır. Şekil 1 de

Y.S.A da kullanılan bir yapay nöron yapısı görülmektedir. Şekildeki X değerleri girdiler olup diğer hücrelerden veya dış ortamdan hücreye giren bilgilerdir. Bilgiler bağlantılar üzerindeki ağırlıklar ile hücreye girer ve ilgili girişin hücre üzerindeki etkisini belirler. Bir ağırlığın değerine göre büyük olması o girişin sinire daha güçlü bağlanması anlamına gelmektedir. Daha sonra bu bilgiler toplama fonksiyonundan geçirilerek bir hücreye gelen net girdiler hesaplanır. Toplama fonksiyonunun değeri her bir ağırlığın ait olduğu girişlerle çarpımının toplamlarına eşik değerinin eklenmesi ile elde edilir. Bu değer transfer veya aktivasyon fonksiyonundan geçirilerek hücre çıktısı elde edilir. Aktivasyon fonksiyonunun sürekli ve türevlenebilir olması ağırlık eğitiminde kullanılan algoritmalar nedeniyle arzu edilmektedir. Literatürde birçok farklı aktivasyon fonksiyonları kullanılmasına rağmen genellikle tanjant-hiperbolik, sigmoid ve doğrusal fonksiyonları tercih edilmekte ve iyi sonuçlar verdiği gözlenmektedir.



Şekil 1. Yapay Nöron

Temel işlemciler olan nöronların birleşmesi ile yapay sinir ağlarının mimarileri oluşturulmaktadır. Şekil 3'de görüldüğü üzere YSA mimarisinde X girdilerin ağı sokulduğu bir giriş tabakası, giriş tabakasından gelen bilgilerin işlenip çıkış katmanına iletildiği bir gizli tabaka ve son olarak da Y çıktılarının elde edildiği bir çıkış tabakası bulunmaktadır. YSA'larının regresyon tekniğine olan temel üstünlerinden biride mimarinin belirlenebilmesi, istenilen sayıda girdi sayısına karşılık, birden fazla çıkış tabakasının bulunabilmesidir.



Şekil 2. Çok Katmanlı YSA Yapısı.

Şekil 2.'de ki YSA mimarisinde 2-5-3 ağ yapısı kullanılmıştır. Bu örnekten anlaşılacağı üzere YSA yı oluşturan nöronlar istenilen sayıda düzenlenebilmektedir. Fakat istenilen sonucun doğru bir şekilde bulunabilmesi için doğru mimariyi bulabilmek önem arz etmektedir. Girdi ve çıktı tabakalarında ki nöron sayıları belli olduğu için çözüm bekleyen sorun gizli katmanda ki nöron sayısının bulunmasıdır. Bu konu hakkında literatürde çeşitli öneriler bulunmasına rağmen kesin bir yöntem bulunmamaktadır. Fakat karmaşık problemlerin çözümünde nöron sayısının artmasıyla daha iyi sonuçlar gözlemlenmektedir.

IV. GENETİK ALGORİTMALAR

Genetik algoritmalar (GA'lar) yapay zeka tekniğinin bir kolu olup, başta mühendislik problemleri olmak üzere bir çok alanda uygulama bulan bir optimizasyon yöntemidir. Bu yöntem, doğadaki canlıların geçirdiği süreç örnek alınarak geliştirilmiş olup, en iyilerin hayatta kalıp kötü nesillerin yok olma ilkesine dayanmaktadır. GA yönteminin esasları ve matematiksel problemlere uygulanması ilk olarak Holland [12] tarafından ortaya konduktan sonra, Goldberg [17], Gen ve Cheng [18], Coley [19] ve Mitchell [20] gibi yazarlar GA tekniğinin ileri kavramlarını ve mühendislikteki uygulamalarını açıklayarak bu yöntemin gelişmesinde etkili olmuşlardır. GA da bir problemin optimum çözümü bazı temel işlemler ile gerçekleştirilir. Bunlar; amaç fonksiyonunun belirlenmesi, kodlama, başlangıç popülasyonunun oluşturulması, üreme, çaprazlama ve mutasyon işlemleridir. GA'nın temelinde popülasyonu oluşturan ve problemin muhtemel çözümünün içeren ve genellikle ikilik sistem ile ifade edilen kromozomlar bulunmaktadır. Kromozomlar semboller zincirinden meydana gelmekte ve her bir sembol bit (digit) olarak adlandırılmaktadır. Kromozomlar ikilik sisteme çevrilerek çaprazlama ve mutasyon işlemlerine uygun hale getirilmektedir. Bit sayısının fazla olması hassaslığın artmasına neden olurken aynı zamanda hesaplama süresinin de uzamasına sebep olmaktadır. Bu nedenle ortalama bir bit sayısının belirlenmesine ihtiyaç duyulur ki bu değer için Denklem 6'nın kullanılması uygun olacaktır.

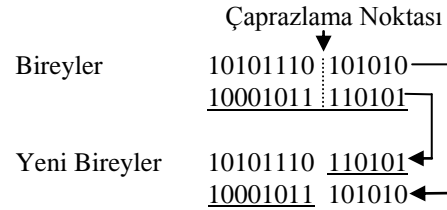
$$2^m \geq \frac{w_i^u - w_i^L}{\Delta w} + 1 \quad (6)$$

Burada; m gerekli olan bit sayısını, w_i^u verilen sınırların üst limitini, w_i^L alt limitini Δw ise hassasiyeti göstermektedir. İkilik sistemdeki sayıların gerçel sayılara dönüşümü ise denklem 7 den yararlanılmaktadır.

$$w_i = w_i^L + \Phi_i \frac{w_i^u - w_i^L}{2^m - 1}, i = 1, 2, 3, \dots, z \quad (7)$$

Burada; Φ_i ikilik sistemden bulunan sayıyı, z popülasyon sayısını göstermektedir.

Döngü esnasında yeni popülasyonların oluşması ve çözüm alanına yaklaşılması gerekmektedir. Üreme, çaprazlama ve mutasyon işlemleri yardımıyla bu süreç gerçekleşmektedir ki bunlar GA'nın üç temel parametresinden biridir. Üreme işlemi, çözüm alanına yaklaşan kromozomları çeşitli teknikleri ile daha güçlü kılarak, bir sonraki popülasyona aktarılma yüzdelarını arttırmaya yönelik bir işlemdir. Üreme işleminde rulet tekeri seçimi, boltzmann seçimi, turnuva seçimi, sıralı seçim kullanılan tekniklerden bazılarıdır. Üreme işleminin tamamlanmasının ardından yeni popülasyonların oluşturulabilmesi için seçilmiş kromozomlar çaprazlanır ve yeni bireyleri oluşması sağlanır. Çaprazlama işleminin bireylerin ne kadarına uygulanması gerektiği çaprazlama yüzdesi (P_c) ile belirlenen bireyler belirlen noktalardan bölünerek karşılıklı olarak yer değiştirir ve yeni bireyler elde edilir [21]. Çaprazlama yüzdesi genellikle 0.4 ile 0.95 arasında değişir. Çaprazlama işleminin bir örneği aşağıda gösterilmektedir.



Mutasyon, oluşan çözümlerin birbirini kopyalamasını önlemek ve sonuca daha hızlı ulaşılmasını sağlayan işlemdir. Örneğin, 1111 yapısında olan kromozom üçüncü bitinde meydana gelen mutasyon ile 1101 kromozomuna dönüşerek işlemi tamamlar. Mutasyon işleminin ne sıklıkla gerçekleştirileceği mutasyon yüzdesi (P_m) ile belirlenir. Yaygın olarak kullanılan mutasyon yüzdesi 0.001 ile 0.01 arasındadır. İstenilen verilerin elde edilebilmesi için GA' da kullanılan amaç fonksiyonunun doğru belirlenmesi gerekmektedir. GA amaç fonksiyonu doğrultusunda en uygun çözüm kümesini bulmaya çalışılmaktadır [21]. Bu çalışmada amaç fonksiyonu, Denklem 8'de görüldüğü üzere ortalama hataların karesi (OHK) olarak biçimlendirilmiştir.

$$OHK = \sum_{i=1}^m (O_{g(i)} - O_{t(i)})^2 \quad (8)$$

Burada; O_g , gerçekleşen değerleri, O_t ise modellerin tahmin ettiği değerleri göstermektedir.

V. İZMİR İLİ İÇİN KAZA TAHMİN MODELLERİ

Trafik kazaları karmaşık olup birçok faktörün bireysel veya birbirleri ile olan etkileşimleri neticesinde meydana gelmektedir. Örneğin, sürücülerin yorgun ve dikkatsiz olması, aşırı hızlı araç kullanmaları, yolun geometrik standartlarının

istenilen niteliklerde olmaması (görüşü sınırlayan engelin bulunması, yetersiz şerit genişliği, banketlerin bulunmaması), araç donanımın eski ve yetersiz olması, atmosferik hava koşullarının sürücü ve araç üzerindeki etkileri vb. kazalara neden olabilmektedir. Ancak kazaya etki eden bütün parametrelerin bir model üzerinde toplanması çoğu zaman mümkün olmamakla birlikte, modelin pratik olarak kullanılabilirliği açısından da uygun değildir. Bundan dolayı geliştirilecek olan modellerin basit ve güvenilir olması arzu edilmektedir. Bu nedenle, İzmir ili için geliştirilen kaza tahmin modellerinde kazalarla doğrudan ilişkili olduğu düşünülen nüfus ve araç sayıları kullanılmıştır. Sadece bu iki parametrenin kullanılmasının diğer nedenleri ise Smeed ve Andreassen kaza modellerinin bu iki parametreyi esas almaları, bu iki parametreye ait verilere kolay ve çabuk ulaşılabilmesi ve geliştirilen modellerinin belli bir güzergâh kesiminden ziyade bütün ili kapsayan genel bir model olmasıdır. Bu çalışmada kullanılan 1986-2005 yıllarını kapsayan nüfus ve araç sayısına ait olan veriler Türkiye İstatistik Kurumundan (T.U.İ.K.)[22], kaza sayıları ise Emniyet Genel Müdürlüğünden (E.G.M) [23] temin edilmiştir.

A. İzmir İline ait Regresyon Analizi (Smeed ve Andreassen) ile Kaza Modelleri

İzmir iline ait kaza tahmin modelleri geliştirilirken, denklem 1 ve 5' de verilen Smeed ve Andreassen model yapılarından yararlanılmıştır. Her iki modelde ölü sayılarını tahmin etmek için önerilmiş olmakla birlikte burada aynı model yapıları kaza sayılarını tahmin etmek için kullanılmıştır. Ancak önceden de belirtildiği gibi Smeed' in yaptığı cebirsel manipülasyon, bağımlı ve bağımsız değişkenleri ve değerlerini değiştireceğinden dolayı burada geliştirilen modelde bağımlı değişken olarak A yerine A/N, N ve P bağımsız değişkenlerinin yerine de N/P kullanılarak regresyon analizi yapılmış ve geliştirilen model denklem 9 da verilmiştir. Denklemde bin araca düşen kaza sayısı verilmektedir. İzmir iline ait geliştirilen bu model şekil yönüyle Smeed modeline benzemekle birlikte başlangıçta alınan bağımlı ve bağımsız değişkenler yönüyle farklılık göstermektedir.

$$\frac{A}{N} = 233,42 \left[\frac{N}{P} \right]^{0,674} \quad (9)$$

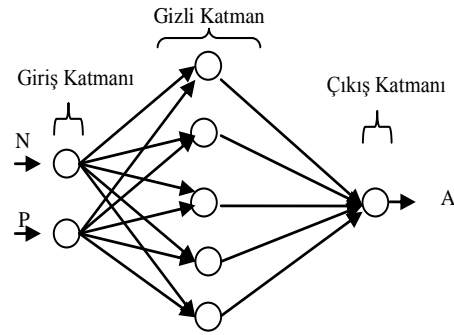
Andreassen model yapısı kullanılarak yapılan regresyon analizi sonucunda ise denklem 10 da verilen model geliştirilmiştir.

$$A = 0,00454 N^{1,574} P^{-0,324} \quad (10)$$

B. İzmir İline ait Yapay Sinir Ağları Yöntemi ile Kaza Modeli

YSA tekniği ile kaza modeli tahmin edilirken birçok farklı ağ mimarisi denenmiş ve Şekil 3'de verilen 2-5-1 ağ mimarisinin bu çalışma için uygun olduğu görülmüştür.

Aktivasyon fonksiyonu olarak birinci katmanda sigmoid, çıkış katmanında ise doğrusal transfer fonksiyonu, öğrenme algoritması olarak da geri yayılım algoritması kullanılmıştır. Şekilde de görüldüğü üzere girdi verileri olarak sadece nüfus ve araç sayısı alınarak kaza sayısı tahmin edilmeye çalışılmıştır. YSA ile yapılan tahminlerde girdi verilerinin sayıları artırılarak daha iyi sonuçların elde edilmesi mümkündür. Ancak bu durumda YSA modeli ile diğer modeller arasında tam bir karşılaştırma yapılamayacağından YSA modelinde de sadece bu iki parametre kullanılarak modellemeye gidilmiştir.



Şekil 3. İzmir İli İçin YSA Kaza Modeline ait Ağ Mimarisi

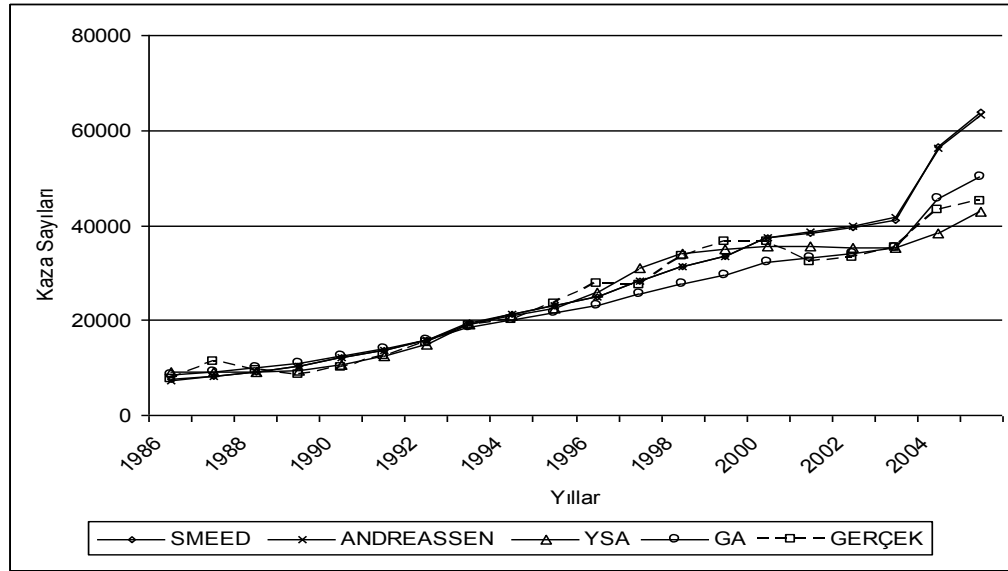
C. İzmir İli İçin Genetik Algoritma Yöntemi ile Kaza Modeli

Genetik algoritma tekniği ile İzmir ili için kaza modeli geliştirilirken diğer modellerde olduğu gibi 1986–2000 yılları arasındaki veriler modelleri geliştirmek için, geriye kalan 5 yıllık veriler ise geliştirilen modelleri test etmek için kullanılmıştır. Literatürde 45 ile 60 arasında önerilen toplam popülasyon büyüklüğü 50, jenerasyon sayısı 200 olarak seçilmiştir. Çaprazlama yüzdesi 0,50, mutasyon yüzdesi ise 0,02 olarak alınmıştır. GA tekniği ile modeller oluşturulurken doğrusal, polinom, üstel ve logaritmik formdaki modeller denenmiş ancak bu çalışma için en başarılı modelin denklem 11' de verilen üstel model olduğu görülmüştür.

$$A = 0,027N^{1,300}P^{-0,214} \quad (11)$$

VI. MODELLERİN KARŞILAŞTIRILMASI VE EN İYİ PERFORMANSA SAHİP MODELİN SEÇİMİ

Farklı teknikler kullanılarak İzmir ili için geliştirilen dört adet kaza modeline ait tahminler ile gerçek kaza değerleri Şekil 4 de verilmektedir. Sonuçlar karşılaştırıldığında, genel olarak dört modelinde ilde meydana gelen kaza artış eğilimini yakaladığı görülmektedir. Ancak bu modellerden geleceğe yönelik tahminlerde hangisinin en iyi performansa sahip olduğu ise istatistikî hata analizleri kullanılarak ortaya konabilmektedir.



Şekil 4. İzmir ili için Smeed, Andreassen, ANN ve GA model tahminleri ve gerçek kaza değerleri

Modellere ait performans ölçüsü olarak denklem12–14’ de verilen ortalama mutlak yüzde hata (OMYH), ortalama karesel hataların karekökü (OKHK) ve ortalama mutlak hata (OMH) kullanılmıştır. Hesaplanan hata değerleri Tablo 1’ de özetlenmekte olup, bu değerler incelendiğinde genel olarak yapay zekâ tekniği kullanılarak geliştirilen modellerin regresyon analizinden üretilen modellere göre çok daha başarılı olduğu görülmektedir. Eğitim periyodunda YSA, test periyodunda da GA kaza modeli seçilen performans ölçütlerinin tümünde minimum hata değerine sahip olmuştur. Smeed ve Andreassen kaza modellerinin performans değerleri eğitim ve test periyotları için birbirine oldukça yakın çıkmıştır. Ancak geleceğe yönelik tahminlerin yapıldığı test periyodunda her iki modelin hata değerleri seçilen hata ölçütlerinin tümünde en yüksek değerlere sahip olmuştur. Bir başka deyişle, bu modellerin test periyodundaki OMYH, OKHK ve OMH değerleri GA kaza modelinin test periyodundaki değerinden sırası ile 5.0, 4.2 ve 5.2 kat daha fazladır.

Genel olarak bu sonuçlar göstermektedir ki, GA kaza modeli en iyi performansa sahip model olup geleceğe yönelik İzmir iline ait kaza tahminlerinde kullanılabilir.

$$OMYH = \frac{1}{n} \sum \left(\left| \frac{o_j - t_j}{o_j} \right| * 100 \right) \quad (12)$$

$$OKHK = \sqrt{\left(\left(\frac{1}{n} \right) \sum_j (o_j - t_j)^2 \right)} \quad (13)$$

$$OMH = \frac{1}{n} \sum (|o_j - t_j|) \quad (14)$$

Tablo 1. Eğitim ve test verileri için kaza modellerine ait hata değerleri

	Smeed		Andreassen		YSA		GA	
	Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test
OMYH	9	20	9	20	7	7	13	4
OKHK	1694	11421	1697	11306	1412	2901	3147	2690
OMH	1389	10186	1373	10198	1094	2474	2395	1968

VII. SONUÇLAR

Bu çalışmada regresyon analizi, YSA ve GA gibi farklı yöntemler kullanılarak İzmir ili için kaza tahmin modelleri geliştirilmiştir. Her üç yöntemde de kaza modelleri geliştirilirken değişken olarak nüfus, araç ve kaza sayıları kullanılmıştır. Regresyon analizi yardımı ile geliştirilen kaza modellerinde Smeed ve Andreassen model yapılarından yararlanılmıştır. YSA kaza modelinde farklı ağ mimarileri denenmiş ve 2-5-1 ağ mimarisinin uygun olduğu görülmüştür. YSA modelinde değişken sayısı artırılarak model performansının yükseltilmesi mümkün olmasına rağmen özellikle girdi verileri olarak sadece araç sayısı ve nüfus kullanılmıştır. Böylece diğer modellerle uyumlu ve eşit koşullar altında performanslarının ölçülmesi sağlanmıştır. Benzer şekilde GA modeli içinde aynı değişkenler kullanılmıştır. Modeller geliştirilirken yalnızca bu değişkenlerin kullanılmasının diğer bir nedeni de bu değişkenlere ait verilerin kolaylıkla elde edilmesidir. Geleneksel ve yapay zeka teknikleri kullanılarak geliştirilen modeller içerisinde en iyi performansa sahip modeli tespit etmek için OMYH, OKHK ve OMH den yararlanılmıştır. Performans analizi sonucunda, Smeed ve Andreassen kaza modellerinin performanslarının her iki periyotta da birbirine yakın olduğu, öte yandan eğitim periyodunda YSA, test periyodunda da GA kaza modelinin en iyi performansa sahip olduğu görülmüştür.

VIII. KAYNAKLAR

- [1] Smeed, R.J. "Some statistics aspects of road safety research", Journal of the Royal Statistical Society, Series A, Part I, 1-34, 1949
- [2] Andreassen., D.C. "Linking deaths with vehicles and population", Traffic Engineering & Control, Vol.26 No.11 pp. 547-549, 1985
- [3] Mekky, A. "Effect of rapid increase in motorization levels on road fatality rates in some rich developing countries" Accident Analysis and Prevention, Vol.17 No.2 pp. 101-109 ,1985
- [4] Chakraborty S, Roy S K. "Traffic accident characteristics of Kolkata" Transport and Communications Bulletin for Asia and the Pacific, No.74, pp.75-86, 2005.
- [5] Valli., P.P. "Road accident models for large metropolitan cities of India" IATSS Research, Vol.29, No.1, pp.57-65, 2005
- [6] Akgungor, A P, Dogan E. "Smeed ve Andreassen kaza modellerinin Türkiye uygulaması: farklı senaryo analizleri" Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, cilt.23, No.4, s. 821-827, 2008.
- [7] Mussone L, Ferrari A, Oneta, M. "An analysis of urban collision using an artificial intelligence model" Accident Analysis and Prevention, Vol.3, No.8, pp.705-718, 1999
- [8] Abdelwahab H T, Abdel-Aty M A. "Development of artificial neural network models to predict driver injury severity in traffic accident at signalized intersection" Transportation Research Record 1746, pp.6-13, 2001
- [9] Delen D, Sharda R, Besson M. "Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks" Accident Analysis and Prevention Vol.38, No.3, pp.434-444, 2006
- [10] Chiou Y.C. "An artificial network-based expert system for appraisal of two-car crash accidents" Accident Analysis and Prevention Vol.38, No.4, pp.777-785, 2006.
- [11] Akgungor A P, Dogan E. "Estimating road accidents of Turkey based on regression analysis and artificial neural network approach" Advances in

- Transportation Studies, An International Journal Section A 16, pp.11-22, 2008.
- [12] Holland, J. H. Adaptation in Natural Artificial Systems, University of Michigan Press, Ann Arbor, Michigan, 211 p. 1992.
- [13] Haldenbilen, S., Ceylan H.. "Genetic algorithm approach to estimate transport energy demand in Turkey" Energy Policy, Vol. 33, Issue 1, pp. 89-98, 2005.
- [14] Gündoğdu, Ö., Gökdağ, M., Yüksel F. "A traffic noise prediction method based on vehicle composition using genetic algorithms" Applied Acoustics Vol.66, No.7, pp. 799-809, 2005.
- [15] Ceylan, H., ve Haldenbilen, S. "Genetik algoritma yaklaşımı ile Avrupa Birliği üyeliği sürecinde Türkiye de beklenen ulaşım talebi ve yönetimi üzerine bir yaklaşım" SDÜ, Fen Bilimleri Enstitüsü Dergisi, cilt. 9, sayı 1, s. 153-159, 2005.
- [16] Akgüngör A. P., Doğan E., "An artificial intelligent approach to traffic accident estimation: Model development and application" Transport, Vol. 24 No. 2, pp.135-142, 2009.
- [17] Goldberg, D. E. Genetic algorithms in search, optimization and machine learning, Addison-Wesley, Harlow, England. 432 p., 1989.
- [18] Gen, M.; Cheng, R. Genetic Algorithms and Engineering Design. New York: John Wiley and Sons. 432 p.,1997.
- [19] Coley, D.A. An introduction to genetic algorithms for scientists and engineers, World Scientific Publishing Company, England 227 p. 1997.
- [20] Mitchell, M. An introduction to genetic algorithms, Cambridge, MA: The MIT Press. 205 p, 1996.
- [21] Elmas, Ç. Yapay zeka uygulamaları, Seçkin Kitabevi, s.425, 2007.
- [22] Türkiye İstatistik Kurumu, Karayolları Kaza İstatistikleri 1986-2005.
- [23] Emniyet Genel Müdürlüğü, Trafik Kaza İstatistikleri, 1986-2005.